AB-47B - Symposium

Career Field Modeling: Estimating Time Utilization in Law Enforcement Patrol Jobs

Charles N. Holt
Jimmy L. Mitchell

Institute for Job & Occupational Analysis
San Antonio TX

John Zuniga
Air Force Security Forces Center
Lackand AFB TX

19990423 046

INTRODUCTION

The Air Force Security Forces Center (AFSFC) was recently established at Lackland Air Force Base, Texas, where it is co-located on the base with the Air Force Military Training Center, which includes the Security Forces Academy among other technical training schools. The AFSFC transferred some Air Staff functions from the Pentagon and other functions from the Air Force Office of Security Police formerly located at Kirtland AFB, NM. The center is responsible for strategic planning, policy, manpower, and budgeting for operation of all Air Force Security Forces operations worldwide.

Given current manpower levels and expected further reductions, the AFSFC has been pondering how to stretch an already overextended manpower force to perform all the tasks and responsibilities for the protection of Air Force personnel, equipment, and resources. A substantial portion of the Security Forces (SF) are already working extended shifts (12 hours) and have limited opportunity for taking annual leave. One suggestion was to reexamine the roles and responsibilities of selected Law Enforcement (LE) jobs, with the view of reengineering the job to reduce nonessential functions. Major Command (majcom) SF staff personnel urged that some objective method be used to systematically collect data on how LE Patrolmen are currently performing their jobs and the time expended on various LE functions.

The AFSFC Plans staff sought the assistance of the Air Force Occupational Measurement Squadron (AFOMS) for descriptive data on LE Patrolmen from the last occupational survey report. They also contacted the Air Force Research Laboratory (AFRL) at Brooks AFB for assistance in collecting actual time data on the tasks performed by LE Patrolmen. Since the study needed to be completed in a very short time frame, it was not feasible to start a new research project; however, since AFRL has an ongoing R&D project to improve survey methodology via experimental studies (GenSurv - see Mitchell, Tucker, Fast, Bennett & Albert, 1997), it was possible to modify one effort to meet both the operational decision making needs of AFSFC as well as the requirements of a technology innovation experiment for GenSurv. Thus these two Air Force requirements could work synergistically.

OPERATIONAL STUDY

In negotiating the requirements of this joint study, AFSFC and AFRL were able to share the workload involved. AFSFC used the AFOMS patrolman job description as a starting point for creating a special LE Patrolman task list. This list was reviewed to identify "core" tasks outlined in AFSF directives as well as other less critical tasks currently being performed by LE Patrolmen.

Methodology

AFRL made the Air Force Survey Authoring System (AFSAS; see Mitchell, Weissmuller, Tucker, Waldroop, & Bennett, 1996) software available and, through the IJOA staff, provided assistance in creating a disk-based automated survey. The AFSFC staff refined the task list, assisted in creating and pilot testing the survey, coordinated with AFOMS to reproduce about 400 disks, selected a representative sample, distributed the surveys with instructions, and monitored returns. With IJOA assistance, AFSFC personnel uploaded data for over 330 diskettes (71% return rate), and provided very detailed quality control of responses. Since these data are for operational decision making, it is imperative that they be as accurate as possible. A number of cases were eliminated if their response patterns suggested they did not use the actual time rating process with reasonable consistency (i.e., some responses were clearly inaccurate such as where their estimate was several times what a

AQI 99-07-1453

qualified subject matter expert thought possible). The removal of such "outliers" (individuals highly divergent from the group mean rating) from the sample is consistent with normal occupational analysis and research practice. It was particularly critical here for both operational and experimental objectives. The final sample consisted of 271 LE Patrolmen from 16 Air Force bases.

Results

Data were summarized using the Statistical Package for Social Sciences (SPSS), version 8.0 PC. A special dos utility was written to calculate individual responses into a common metric - hours per task per year. SPSS was used in lieu of CODAP, since the normal occupational analysis software cannot handle multiple digit ratings (normal 1 to 9 relative time spent ratings are single digit). Actual time spent data were summarized for "core tasks" versus those tasks not as critical to LE patrol functions; the data were displayed as "percentage of work time" which could then be used with manpower standards to calculate possible savings.
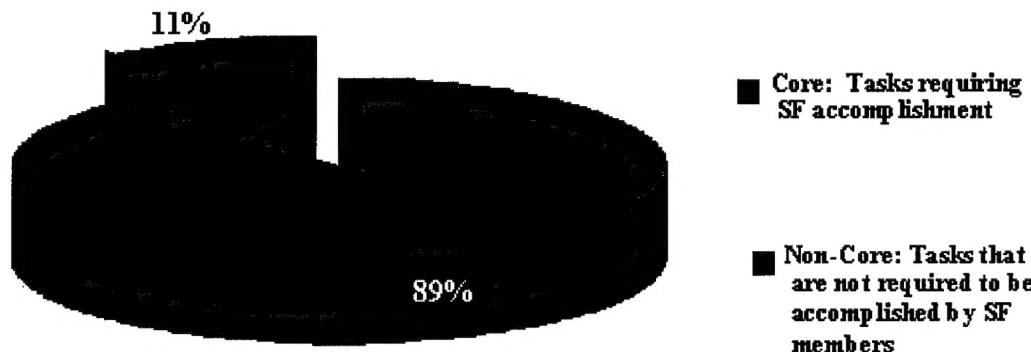

Figure 1. - Core Versus Non-Core Patrol Tasks

This analysis clearly indicated that there were substantial savings to be made by eliminating some of the non-core tasks; that is by changing how such functions are accomplished. Similar data were displayed for the various major commands demonstrating where substantial efficiencies could be achieved. Actual manhour savings were quantified and evaluated.

The AFSFC staff developed a number of possible policy options which could be made to implement this job reengineering, as well as the relative manhour impacts for each. These options were briefed to senior AFSFC executives (General Officer level). Such policy changes include transferring responsibility for minor incident investigation and reporting to the desk sergeant (individuals will report to the desk), escorting only Air Force funds during transfer (banks, etc. will provide their own escorts), and eliminating some tasks. Some of these proposed changes were approved and additional options are now being staffed.

Demographic data for the sample were also summarized and briefed to demonstrate the typical working conditions (shifts worked, hours per day, annual leave taken, etc.) as well as job attitudes and career intentions. Such data indicate that over half the force is working extended 12-shifts and many work six day weeks, but a majority have generally positive attitudes toward their job, use of their talents and training, and expectation of an Air Force career. A typical workweek for patrolmen is about 56+ hours versus the traditional 40 hours.

Overall, this operational study was an outstanding success. Quantitative actual time estimates were collected in automated form very quickly, were compiled and summarized, then synthesized for various options to provide specific manyear implications for each. The data were used to support possible policy changes and senior executives made appropriate decisions. All this was completed just four and a half months after the first AFSFC meeting at AFRL.

EXPERIMENTAL STUDY

The primary objective of the experimental phase of this study was to determine if feedback, in the form of a continuously accruing total-time-accounted-for display, could help to improve the reliability and validity of actual time spent data. This was to be assessed by contrasting two groups, one which received feedback and one with no such feedback.

Methodology

Data collection of actual time spent per task information has been successfully collected in earlier proctored field experiments by Air Force Research Laboratory scientists (Albert, Phalen, Selander, Dittmar, Tucker & Weissmuller, 1994) using software installed on a personal-size computer (PC). Actual time data can be a superior metric for many purposes, in that it has almost unlimited variance and can be used to compare across jobs, occupations, organizations, etc (Phalen, 1995). A modified form of the actual time software (to fit on a high density diskette as opposed to operating from a PC hard disk) has been used in collecting actual time data with Basic Military Training Instructors (Albert, Bennett, Pemberton, Holt & Waldroop, 1997) and is the software used for this study. Two separate forms of the survey were produced, one of which had specially-developed software to display a running total of hours and percentage of time accounted for; the second form had this display disabled. Disks were reproduced for both forms, which were equally distributed to each of the sixteen bases surveyed; thus an incumbent at a base had an equal chance of receiving of either version.

Since all survey participants belong to the same job type (LE Patrol), the normal occupational analysis variations of jobs within the occupation were eliminated. Thus, the present study reduces many of the usual sources of variance in occupational data and such extra variance should not be a problem here. There are, however, some other expected types of variance involved, particularly the major differences in actual hours between the "normal" Air Force work schedule (8-hour shifts) versus the "extended" (12-hour) shift work now required at many bases. The sample was selected to insure that bases on these various schedules were included systematically. Some analysis needs to be done to highlight the differences in actual time estimates for various shift options.

Within both surveys, the task list was organized into major duties of the patrolman jobs. The incumbent was asked to rate the importance of each duty to the job, and the software then administered the survey in descending order of rated importance. This new administration technique to some degree controls for rater fatigue by insuring that major duties of the job are considered first and other tasks are rated later. Recently, this technique has also been used successfully with a 20,000 case study for another service. In that study, the software also screened by skill-level so that only those tasks appropriate to the individual's skill level were rated. Data collected were processed to yield "hours per year" as a common metric for the this analysis. Data analysis including testing between-group differences in mean and standard deviation was accomplished using the Statistical Package for the Social Sciences (SPSS) employing traditional t-tests.

Results

The major contrasts to be made between the "feedback" and "no feedback" groups involve the means and standard deviations of the two groups. It was anticipated that there be no mean difference between the groups, but that the "feedback" group would have a smaller standard deviation if the feedback actually had an impact on the estimates the individuals were making. All ratings were averaged across individuals and then across all tasks and SPSS used to calculate group statistics (see Table 1).

Table 1 - Pair Sample Statistics for Feedback versus No Feedback Conditions

Statistic by Group	Mean	N	Standard Deviation	Std. Error Mean	Correlation
Mean -Feedback	59.826	114	81.843	7.665	.748 *
No Feedback	63.638	114	84.710	7.934	
S.D. - Feedback	107.198	114	137.746	12.901	.732 *
No Feedback	116.065	114	151.379	14.178	

* Significant at p. > 001

Note that the means and standard deviations are both higher for the No Feedback group than for the Feedback

group. There is a high correlation between the task averages for the two groups, as would be expected. It is also worth observing that the standard deviations are all higher than the mean and standard deviations for both groups; this finding suggest that there are considerable sources of variation in the ratings (between tasks and among individual raters) not associated with the experimental condition (feedback or no feedback status). The trend to smaller standard deviation for the feedback group was expected, but the trend to a lower mean was not. The next question is, of course, is whether the differences in group means and standard deviations are statistically significant. The following t-test was performed to address this issue.

Table 2 - Paired Comparison of Feedback versus No Feedback Groups

	Paired Differences			t
	Mean Difference	**Std. Dev.**	**Std. Error Mean**	
Mean Feedback - Mean No Feedback	- 3.81288	59.226410	5.547062	- .687
S.D. Feedback - S.D. No Feedback	- .86662	106.649243	9.988617	- .888

These values indicate that there are no statistically significant differences in either the mean or the standard deviation between these groups. Thus, even with the trend toward less standard deviation for the no feedback group, the difference is not great enough to prove the effect. This lack of significance may result from the high standard deviations in all of the ratings which was noted above. This may be, in part, a function of having included both 8-hour shift workers with 12-hour shift personnel in both types of groups; obviously those working 12-hour shifts will have greater numbers of hours per year worked for most tasks. If this factor is a primary complicating factor here, then perhaps we should do our analysis with the feedback and no feedback data subdivided by what shift incumbents are working.

Data were resorted and a new analysis undertaken to assess this potential sources of variance. In this analysis, the data represent total number of hours per year worked, summed across all tasks and averaged across individuals. The sample was relatively balanced with 60 individuals in the 8-hour shift feedback group, and 76 in the 12-hour feedback group. For the no feedback group, 58 individuals worked the 8-hour shift and 76 were in the 12-hour shift group (total of 270 cases). Results of the by-shift versus feedback group analysis are shown in Table 3 below.

Table 3 - Analysis of Variance Feedback Condition and Shift Hours Worked

	Sum of Squares	**df**	**Mean Square**	**F value**	**Significance**
Between Groups	1702912.435	3	567637.478	.068	.977 *
Within Groups	2205019752.860	266	8289547.943		
Total	2206722665.295	269			

* Not Significant

This total lack of a statistically significant result clearly demonstrates that the difference in shift, although great,

4

is not the primary causal factor. Rather, the individual differences are so large that they overwhelm all other sources of variance and prevent any trends in the data from even approaching significance.

DISCUSSION

One major reason for collecting actual time estimates is to increase the variance in ratings in order to overcome the restriction in range problem with other types of rating scales (Phalen, 1995). A number of researchers have maintained that "higher variability in item responses is indicative of higher data quality (Stanton, 1998, page 713)." Clearly the present study was successful in developing considerable more variability that would have been possible using the normal "relative time spent" scale used in most occupational analysis studies. Part of the variance in ratings was a function of the current unusual work schedules of the majority of Air Force Law Enforcement Patrolmen; over 50 percent of the members in the job are working in excess of 12 hours per day and many are working six days a week. Reengineering the job to reduce many of the non-critical functions is an extremely worthwhile objective, and the fact that this study helped to generate and justify executive decisions in that direction is an excellent outcome.

The results of the experimental study, while they did not fully conform to our expected results, tended to be in the direction anticipated in that there was some reduction in the standard deviation for the feedback group versus the no feedback group (albeit not a statistically significant result). Further analysis of the data suggests that there is some excess variability in the ratings and perhaps some overestimate of the amount of work time for some respondents. Examination of individual responses revealed that some respondents were not using a consistent frame of reference when rating individual tasks and appeared to be estimating actual time to perform the task inappropriately. While the more extreme cases could be identified and eliminated as outliers, eliminating too large a portion of your sample this way would border on selecting your data to fit your expected conclusion.

Another possible problem is whether the tasks to be rated are well written, reasonably discrete, and time rateable as recommended by most experienced occupational analysts (Archer & Fruchter, 1963; Christal, 1974; Driskill & Gentner, 1978). If the tasks in a job inventory are not mutually exclusive or tend to be ambiguous, the ratings will tend to be more diverse but possibly spurious, and the result will be an overestimate of the time spent on a given task or function; likewise total hours worked would be exaggerated. Review of the task list for this study indicates there may have been some lack of discreteness for a few tasks, particularly when some are somewhat global statements (i.e., patrol the base, etc.). Overall, it was a fairly good task list but if the study were ever repeated, some additional polishing of the task list with the traditional task writing criteria in mind would be worthwhile as well as extensive subject-matter expert review.

Another factor which may have introduced extra variance in responses was the lack of some of the proctoring of responses which was part of the original laboratory study (Albert, et al., 1994; Phalen, 1995). In this hard disk software, certain screening criteria were built in so that if a response was extreme (i.e., exceeded the maximum expected level) then the software put up an alert flag which asked the respondent to reconsider his or her rating (Ibid). When the software was simplified for the field feasibility study (Mitchell, Weissmuller, Bennett, Agee, & Albert, 1995) so that it could be exported on low density diskettes to Air Force worldwide locations (and run from disk without installing on the PC hard disk), the extra monitoring of responses and prompting raters to reconsider had to be eliminated. Clearly, such software proctoring would be worthwhile in helping to keep down overestimation and making rater responses more realistic. Such software proctoring would be much easier to implement in a Windows environment than is possible with the current DOS-based system (OASurv).

Further research and development to operationalize actual time spent data collection would certainly be worthwhile. While this study was extremely successful in meeting the need for quick actual time data for a selected job, the experimental phase of the project was not totally successful. It would be very worthwhile to recollect such data, once a Windows version of the software is available or when the internet enabled product (GenSurv) becomes fully operational. If the GenSurv system is to be used to collect actual time spent data in addition to the traditional relative time spent and training evaluation data, then the system should be modified to include response monitoring and real time prompting of raters to reconsider extreme ratings. In addition, if the trends found in the current study toward lower standard deviation of responses when feedback of time accounted for is provided can be verified through additional studies with improved software, then GenSurv should probably also include the running total time functionality.

References:

Albert, W.G., Bennett, W., Jr., Pemberton, K., Holt, C.N., & Waldroop, P. (1997). Evaluation of automated technology in basic military training using computer-based surveys. Presentation in the

symposium, J.S. Tartell and H.W. Ruck, co-chairs, Advanced Technology Research and Applications in Occupational & Training Analysis & Organizational Assessment, at the 38th annual conference of the International Military Testing Association (IMTA), Sydney, Australia (*Proceedings* in press).

Albert, W.G., Phalen, W.J., Selander, D.M., Dittmar, M.J., Tucker, D.L., & Weissmuller, J.J. (1994). Large-scale laboratory test of occupational survey software and scaling procedures. *Proceedings of the 36th Annual Meeting of the International Military Testing Association*. Rotterdam, The Netherlands: European Members of the IMTA.

Archer, W.B. & Fruchter, D.A. (1963). *The construction, review, and administration of Air Force job inventories* (PRL-TDR-63-21). Lackland AFB, TX: 6570th Personnel Research Laboratory.

Christal, R.E. (1974). Collecting, analyzing, and reporting information describing jobs and occupations. (AFHRL-TR-74-19, AD-774 575). Lackland AFB, TX: Occupational Research Division, Air Force Human Resources Laboratory.

Driskill, W.E., & Gentner, F.C. (1978). *Four fundamental criteria for describing the tasks of an occupational specialty* (in Technical Note 78-04). U.S. Air Force Occupational Measurement Center, Randolph AFB, TX.

Mitchell, J.L., Tucker, D., Fast, J., Bennett, W., Jr., & Albert, W.G. (1997, October). Research and development of new occupational analysis and training evaluation technologies. Presentation in the symposium, J.S. Tartell and H.W. Ruck, co-chairs, Advanced Technology Research and Applications in Occupational & Training Analysis & Organizational Assessment, at the 38th annual conference of the International Military Testing Association (IMTA), Sydney, Australia (*Proceedings* in press).

Mitchell, J.L., Weissmuller, J.J., Bennett, W., Jr., Agee, R.C., & Albert, W.G. (1995). Final results of a field study of the feasibility of computer-assisted occupational surveys: Stability of task and job information. *Proceedings of the 37th annual conference of the International Military Testing Association (IMTA)*, pp. 231-236. Toronto, Ontario, Canada: Canadian Forces Applied Research Unit.

Mitchell, J.L., Weissmuller, J.J., Tucker, D.L., Waldroop, P., & Bennett, W., Jr. (1996, November). Development and application of a computer-assisted survey authoring tool for training needs assessment. In the symposium, H. W. Ruck, Chair, Recent Research and Applications in Training Needs Assessment and Evaluation. *Proceedings of the 38th Annual Conference of the International Military Testing Association*, pp.486-491. San Antonio, TX: Air Force Personnel Center, Armstrong Laboratory Human Resources Directorate, & the Air Force Occupational Measurement Squadron.

Phalen, W.J. (1995). A critical evaluation of various procedures for estimating time spent. *Proceedings of the 37th Annual Conference of the International Military Testing Association,* pp. 418 - 423. Toronto, Ontario, Canada: Canadian Forces Applied Research Unit.

Stanton, J.M., (1998). An empirical assessment of data collection using the internet. *Personnel Psychology*, **51**:709-725.

INTERNET DOCUMENT INFORMATION FORM

A . Report Title: Career Field Modeling: Estimating Time Utilization
in Law Enforcement Patrol Jobs

B. DATE Report Downloaded From the Internet   4/21/99

**C. Report's Point of Contact: (Name, Organization, Address,
Office Symbol, & Ph #):   Navy Education and Training
Professional Development and
Technology Center
Navy Advancement Center Dept
· Dr. Grover Diesel, (850) 452-1815
6490 Saufley Field Road
Pensacola, FL 32509-5237**

D. Currently Applicable Classification Level: Unclassified

E. Distribution Statement A: Approved for Public Release

**F. The foregoing information was compiled and provided by:
DTIC-OCA, Initials**: VM_ **Preparation Date**: 4/22/99__

The foregoing information should exactly correspond to the Title, Report Number, and the Date on
the accompanying report document. If there are mismatches, or other questions, contact the
above OCA Representative for resolution.